

Simplified K -RF Distances for Labeled Directed Acyclic Graphs

- “sK-RFmeasures.py” contains the code to compute pairwise simplified K -RF distances of a set of labeled DAGs stored in an input JSON file. The command line for running the code is “python3 sK-RFmeasures.py inputfile k”, where “inputfile” refers to a JSON file that consists of the DAGs whose pairwise distances need to be computed, and k is an integer that is greater than or equal to zero.

A labeled DAG G in the input file needs to be represented as $[A, B]$, where A is the list of nodes in G with their labels, and B is the list of edges of G . Each node $v \in V(G)$ is represented as $[v, a_1, \dots, a_n]$ if $\ell(v) = \{a_1, \dots, a_n\}$. Note that v must be in the first position of the array. Additionally, each edge from $v \in V(G)$ to $u \in V(G)$ is represented as $[v, u]$.

- “sKfunctions.py” contains the necessary functions required to run the code in “sK-RFmeasures.py”.
- “255DAGs.json” is a sample input file containing all 255 DAGs used for our clustering experiment described in the paper.
- “250DAGsPC.json” is a sample input file containing all 250 DAGs used for the correlation analysis of the simplified K -RF measures with generalized Kendall-tau and Katz similarity distances, as described in the paper.
- “1000one-label-trees.json” and “976one-label-trees.json” are another sample input files containing the datasets used for the correlation analysis of the simplified K -RF measures with $CASet\cap$, $DISC\cap$, and GRF , as described in the paper.
- “1000one-label-trees.json” and “976one-label-trees.json” are sample input files containing the datasets used for the correlation analysis of the simplified K -RF measures with $CASet\cap$, $DISC\cap$, and GRF , as described in the paper.